METHODS

Multiplex PCR-Based Real-Time Invader Assay (mPCR-RETINA): A Novel SNP-Based Method for Detecting Allelic Asymmetries Within Copy Number Variation Regions

Naoya Hosono,[1]* Michiaki Kubo,[1] Yumiko Tsuchiya,[1] Hiroko Sato,[1] Takuya Kitamoto,[1] Susumu Saito,[1] Yozo Ohnishi,[2] and Yusuke Nakamura[1,2]

[1]Laboratory for Genotyping, SNP Research Center, The Institute of Physical and Chemical Research (RIKEN), Kanagawa, Japan; [2]Laboratory of Molecular Medicine, Human Genome Center, The Institute of Medical Science, University of Tokyo, Tokyo, Japan

Communicated by Michael Dean

We report the development of a real-time Invader assay combined with multiplex PCR (mPCR-RETINA), an SNP-based approach that can measure the allelic ratio in copy number variation (CNV) regions of a genome. RETINA monitors the real-time fluorescence intensity of each allele during the Invader assay and detects allelic asymmetries caused by genomic duplication/multiplication in heterozygous individuals. By combining mPCR-RETINA and real-time quantitative PCR that detects total copy number, we can estimate the copy number of each allele in CNV regions, which should be useful for investigating the functional significance of allele copy number with disease susceptibilities and drug responses. Also, mPCR-RETINA can efficiently refine the detailed structures of CNV regions. Due to the combination of RETINA with multiplex PCR, mPCR-RETINA requires a very small amount of genomic DNA for analysis (0.1–0.38 ng/locus). Additionally, mPCR-RETINA has clear advantages in its simple protocol and target-specific reaction, even in nonunique regions. We believe mPCR-RETINA will provide a significant contribution to identifying functional alleles in CNV regions. Hum Mutat 0, 1–8, 2007

KEY WORDS: CNV; duplication; multiplication; SNP; multiplex PCR; real-time Invader assay

INTRODUCTION

Copy number variations (CNVs) are relatively common in the human genome, estimated to be approximately 12% of the all genomic regions [Redon et al., 2006]. Since copy number differences in functional genes may affect the quantity of gene products, some of them are likely to be associated with susceptibility to various diseases as well as the efficacy or adverse reaction to certain drugs [Gasche et al., 2004; Gonzalez et al., 2005; Aitman et al., 2006; Padiath et al., 2006]. For example, the multiplication of functional alleles in the *CYP2D6* gene was shown to be correlated with ultrarapid metabolization of codeine, which results in life-threatening opioid intoxication [Gasche et al., 2004]. Since qualitative (allele copy number) and quantitative (gene copy number) alterations affect critical phenotypes, it is essential to develop new methods that can measure copy number values of both genes and alleles to investigate the functional significance of CNVs [Freeman et al., 2006].

So far, more than 2,000 CNV regions have been identified by various methods, including BAC-array comparative genomic hybridization (CGH) [Redon et al., 2006; Iafrate et al., 2004; Sharp et al., 2005; Locke et al., 2006], an oligonucleotide array called ROMA [Sebat et al., 2004], fosmid paired-end sequence mapping [Tuzun et al., 2005], and the SNP mapping array [Redon et al., 2006], and are summarized in the Database of Genomic Variants (http://projects.tcag.ca/variation). However, the genomic regions that are deleted or duplicated/multiplied are not well defined due to the technical limitations in these methods: for example, BAC-array CGH detects relatively large CNV regions (>50 kb); the resolution in the ROMA method is low due to its low coverage of the genome; the sample size used for fosmid paired-end sequence is very small; the marker density in the SNP mapping array is low in some parts of the genomic regions [Sharp et al., 2006; Freeman et al., 2006]. Thus, many of the reported CNV regions must have their detailed structures defined to investigate the presence or absence of a copy number difference in the functional gene unit. Recently, high-density oligonucleotide tiling array CGH has been frequently used to precisely define CNV breakpoints [Locke et al., 2006; Urban et al., 2006]. This method

The Supplementary Material referred to in this article can be accessed at http://www.interscience.wiley.com/jpages/1059-7794/suppmat.

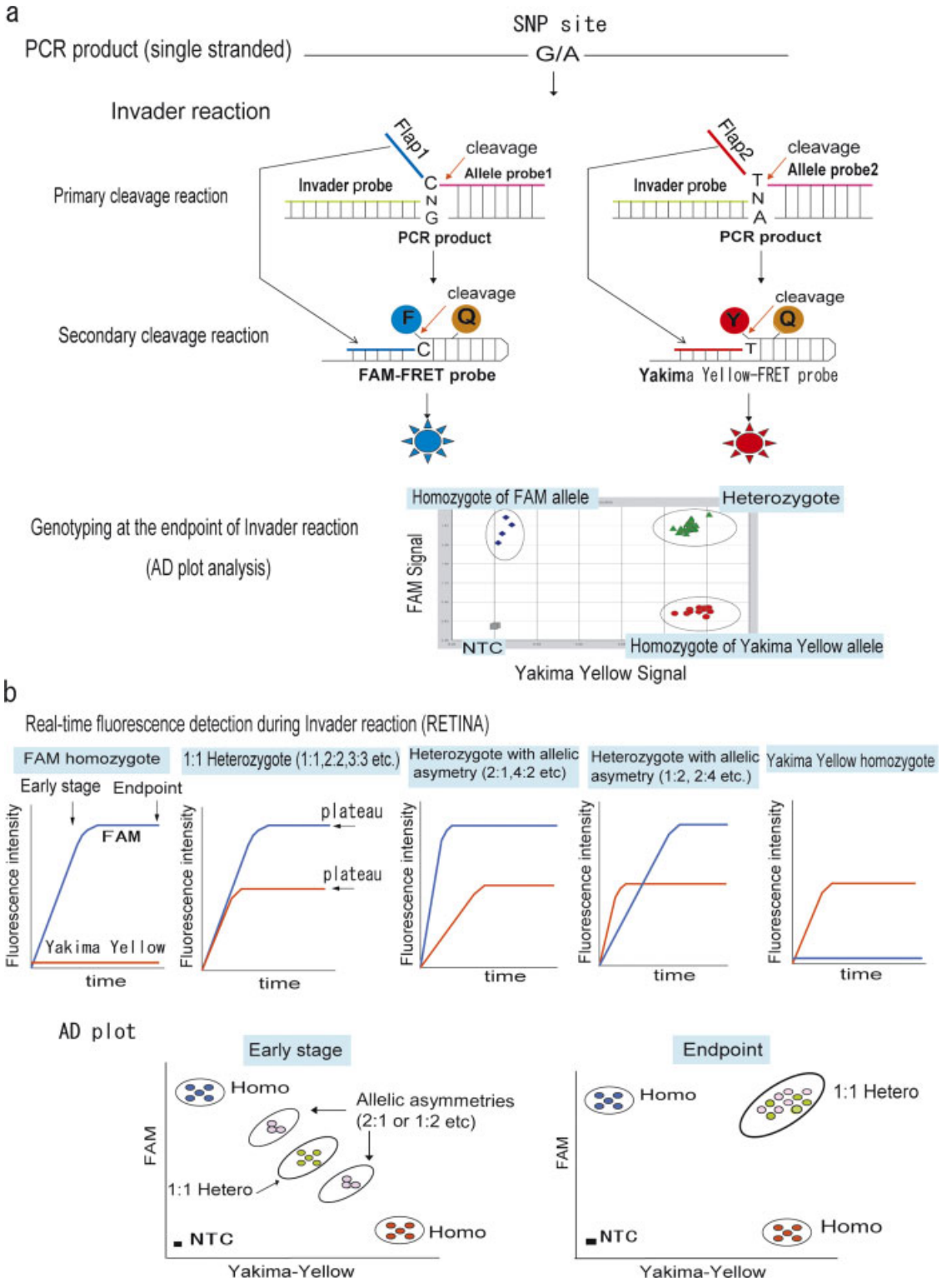
Received 9 March 2007; accepted revised manuscript 14 June 2007.

*Correspondence to: Naoya Hosono, 1-7-22 Suehiro-cho, Tsurumi, Yokohama, Kanagawa, 230-0045, Japan.
E-mail: nhosono@src.riken.jp

Grant sponsor: Ministry of Education, Culture, Sports, Science and Technology.
DOI 10.1002//humu.20609
Published online in Wiley InterScience (www.interscience.wiley.com).

has the advantage of high resolution, but its capability in the target-specific recognition of probes is not sufficient, especially in nonunique regions that are known as CNV hotspots [Locke et al., 2006; Sharp et al., 2006].

The Invader assay, coupled with multiplex PCR, is one of the SNP genotyping methods with the highest accuracy [Ohnishi et al., 2001; International HapMap Consortium, 2005]. The basic principle of the PCR Invader assay is shown in Fig. 1a. In our routine SNP genotyping, we simultaneously amplified 96 fragments of SNP sites by multiplex PCR. After that, the PCR products were put into 384-well plates, and the Invader assay was performed for each SNP with unique probes corresponding to each allele [Ohnishi et al., 2001]. This method was originally developed as an endpoint assay, and fluorescence intensities were only measured after a 15-minute to 60-minute incubation of the Invader reaction [Ohnishi et al., 2001; Nevilie et al., 2002]. Although this method



generates clear and accurate genotyping results, it was not investigated as to whether this assay could be applied to the detection of CNVs. In this study, we modified the Invader assay to detect real-time fluorescence during the enzymatic reaction after multiplex PCR (multiplex PCR real-time Invader assay [mPCR-RETINA]). We show that mPCR-RETINA can measure the allelic ratio utilizing allelic asymmetries caused by genomic duplication/multiplication in heterozygous individuals, and determine the allele copy number in CNV regions by combining the total copy number data obtained from real-time quantitative PCR. We also show that mPCR-RETINA can be useful for refining breakpoints within duplicated/multiplied regions.

MATERIALS AND METHODS

Genomic DNA

We used 180 genomic DNA samples consisting of 30 trios of European ancestry (CEU) and 30 trios of Yoruba (YRI) that were for the International HapMap project [International HapMap Consortium, 2005]. These DNA samples were purchased from the Coriell Cell Repositories (Camden, NJ).

PCR-RETINA

We used the reported PCR primers, invader probes, and allele probes for the *CYP2D6* assays [Nevilie et al., 2002]. The remaining PCR primers were designed by Primer Express 1.5 (Applied Biosystems, Foster City, CA), and the other invader probes and allele probes were designed and synthesized under the reported criteria [Mast and de Arruda, 2006]. The sequences of all primers and probes are listed in Supplementary Table S1 (available online at http://www.interscience.wiley.com/jpages/1059-7794/suppmat). Fluorescence resonance energy transfer (FRET) probes labeled with FAM or Yakima Yellow were purchased from Third Wave Technologies (Madison, WI). Rox dye (Sigma, St. Louis, MO) was used for the normalization of reporter signals. We used Takara Ex Taq HS (Takara, Shiga, Japan) for all PCR amplification according to the manufacturer's instructions with a primer concentration of 100 nM in all assays. PCR was performed on GeneAmp 9700 (Applied Biosystems) with a reaction volume of 5 μl in a single PCR. The PCR condition of the *CYP2D6* assays was initiated at 95°C for 2 minutes followed by 35 cycles at 95°C for 15 sec and 68°C for 4 minutes. The PCR condition of the *MRGPRX1* assays was initiated at 95°C for 2 minutes followed by 35 cycles at 95°C for 15 sec, 58°C for 30 sec, and 72°C for 1 minute. For the multiplex PCR, the PCR condition was initiated at 95°C for 2 minutes followed by 37 cycles at 95°C for 15 sec, 58°C for 30 sec, and 72°C for 4 minutes with a reaction volume of 10 μl [Ohnishi et al., 2001]. After PCR, the products were diluted up to 10-fold and used as templates for RETINA. We carried out RETINA for each SNP with the reaction volume of 4 μl on ABI prism 7900 (Applied Biosystems) following the protocol recommended by Third Wave Technologies. Data analysis was performed with Excel (Microsoft, Redmond, WA).

Estimating Total Copy Number by Taqman Assays

We carried out Taqman assays to estimate the total copy number of CNV regions. We first used the reported Taqman (Applied Biosystems) assay for *CYP2D6* [Bodin et al., 2005]. However, since we found a three-base insertion in some YRI individuals at the reverse primer site of the reported assay, we designed a new reverse primer by Primer Express and then performed copy number analysis again to obtain accurate data in the analysis of YRI individuals. For *MRGPRX1*, we designed all assays using Primer Express 1.5. These Taqman probes were labeled with FAM at the 5′ end and linked by nonfluorescence quencher (NFQ) and minor groove binder (MGB) at the 3′ end. As the reference gene, we used the RNase P assay (Applied Biosystems) labeled with VIC. All Taqman assays were performed following the reported protocols and copy number calculation was conducted by the delta-delta threshold cycle (Ct) method [Bodin et al., 2005]. We assumed the samples with a median delta Ct value were two-copy and used them as a calibrator. All samples were examined in duplicate and the average copy number values were used in the scatter plot analysis. The primer and probe sequences of all assays are listed in Supplementary Table S1.

Direct Sequencing

Two YRI samples showing discordant results between PCR-RETINA and the Taqman assay were amplified by PCR under the reported reaction condition [Dorado et al., 2005]. The amplified DNAs were subjected to direct DNA sequencing on the ABI Prism 3700 sequencer (Applied Biosystems) and analyzed with Polyphred software (University of Washington, Seattle, WA; http://droog.gs.washington.edu/PolyPhred.html). The primer sequences for direct sequencing are listed in Supplementary Table S1.

RESULTS

Detection of Allelic Asymmetries Using Artificial Templates

We first synthesized two artificial 80-base oligonucleotide templates corresponding to the genomic sequence of an

FIGURE 1. **a: A schema of SNP typing by the PCR Invader assay. The PCR Invader assay is a simultaneous detection method for two different alleles in an SNP. Two allele probes are designed for each target SNP, complementary to each allele. Each allele probe has a different universal flap sequence at the 5′ end. One invader probe is designed at the opposite site of allele probes. In the primary reaction, Cleavase VIII cuts the allele probes with a single-base invasive structure by an invader probe at the SNP site when the allele probes are hybridized to the complementary target DNA. The red arrows indicate the cleavage sites. The cleaved allele probes with a flap and a nucleotide of the SNP site from the primary reaction hybridize to the specific FRET probes and behave as invader probes in the secondary reaction. Cleavase VIII cuts the FRET probe with an invasive structure by the flap sequence, resulting in fluorescence signal generation. The primary and secondary Invader reactions occur simultaneously. If the probes do not hybridize perfectly at the site of interest and no overlapping structure is formed, no cleavage occurs and no fluorescence signal is generated. SNP genotyping is usually performed by clustering analysis in a two-dimensional allelic discrimination (AD) plot at the endpoint of the reaction. b: The principle of detecting allelic asymmetries by RETINA. In the PCR Invader assay, fluorescence signals are usually saturated and reach the plateau phase within 5 to 10 minutes (early stage) of the Invader reaction. This plateau effect induces all samples with various allelic asymmetries to a 1:1 fluorescence intensity ratio and allelic asymmetries cannot be discriminated in the AD plot at the end point of the reaction. Samples with allelic asymmetries show different patterns from a 1:1 heterozygote in real-time fluorescence signal curves and RETINA can discriminate these samples in the AD plot at an early stage of the reaction (before saturation). The time point just before the reaction is saturated usually provides the best separation of the clusters with allelic asymmetries.**

rs2114912 SNP locus to investigate the detection capability of allelic asymmetries in the PCR Invader assay (the sequences are shown in Supplementary Table S1). We generated standard samples with a range of allelic ratios from 8:1 to 1:8 using these oligonucleotides. After the amplification of DNA fragments by PCR, we performed real-time fluorescence detection every 30 seconds during the 30-minute Invader reaction (RETINA). In allelic discrimination (AD) plot analysis, the samples with various allelic asymmetries in the concentration of each allele were clearly separated in proportion to the allelic ratio in the early stage of the reaction, but were merged as one heterozygote cluster in 20 minutes or more of the reaction (Fig. 2b–d). Among the AD plots

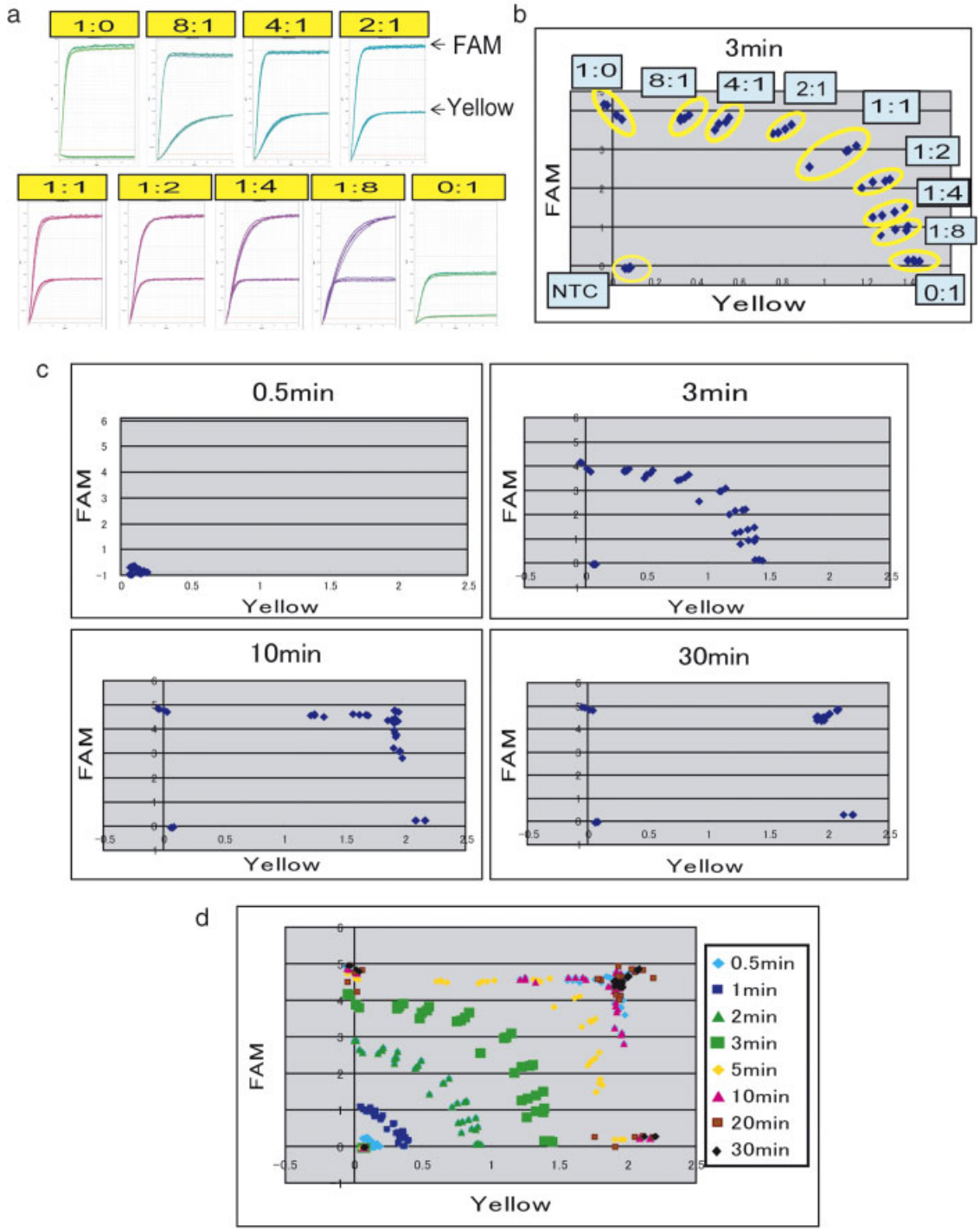


FIGURE 2. **Detection of allelic asymmetries by PCR-RETINA using artificial templates. All experiments were performed in quadruplicate. a: Real-time detection of fluorescence signals during the Invader assay. The ratios in the boxes indicate the allelic ratio of each standard sample. Reporter fluorescence signal values (FAM and Yakima Yellow) after the normalization using passive reference (Rox) were plotted. The vertical axis is the normalized reporter signal (FAM or Yakima Yellow), and the horizontal axis is the reaction time every 30 seconds. b: The AD plot of PCR-RETINA at 3 minutes in the Invader reaction. The vertical axis is the normalized FAM allele signal, and the horizontal axis is the normalized Yakima Yellow allele signal. c: The AD plots at 0.5, 3, 10, and 30 minutes in the Invader reaction. d: Transition of cluster patterns showing allelic asymmetries during the 30-minute Invader reaction.**

at various time points during the Invader reaction, the time point just before the FAM or Yakima Yellow fluorescence signal became saturated could provide the best separation of the clusters with allelic asymmetries. The biochemical schema of PCR-RETINA is shown in Fig. 1b.

Detection of Gene Duplications/Multiplications in the Human Genome

To examine whether PCR-RETINA could be applied to allelic-asymmetry detection within CNV regions, we investigated two representative CNVs corresponding to the *CYP2D6* and *MRGPRX1* loci [Lovlie et al., 1996; Tuzun et al., 2005; Redon et al., 2006]. We selected three SNPs for *CYP2D6* from a previous report [Nevilie et al., 2002] and three SNPs for *MRGPRX1* from the dbSNP database (the SNP locations are shown in Supplementary Figs. S1a and S2a, respectively). In the *CYP2D6* assay, we found that two CEU subjects and nine YRI subjects were placed outside of the three major clusters in the AD plots in at least one of the three loci (Supplementary Fig. S1b and c). We confirmed all of these 11 samples to have three copies of *CYP2D6* by means of the reported Taqman assay (Supplementary Fig. S1d and e) [Bodin et al., 2005]. On the other hand, PCR-RETINA was unable to identify some individuals who were indicated to have three copies or one copy by the Taqman assay, because they contained three copies of the same allele or a deletion of one allele.

PCR-RETINA analysis at the *MRGPRX1* locus revealed several clusters in the AD plots of both populations. Individuals placed outside of the major clusters were confirmed to have three or more copies by the Taqman assay (Supplementary Fig. S2b–e). In this evaluation, PCR-RETINA was also unable to identify a few individuals having four copies, because they had two copies each of both alleles and were plotted in the heterozygote cluster. These experiments indicate that PCR-RETINA could detect individuals with allelic asymmetries, but could not detect individuals with multiplied allelic symmetries, homozygous individuals with duplications/multiplications, or individuals with deletions.

In the experiment of *CYP2D6* using YRI samples, we observed discordances in the results of two samples genotyped by PCR-RETINA and the reported Taqman assay. These samples were estimated to have one copy (a deletion in one chromosome) in the reported Taqman assay, but PCR-RETINA judged these two individuals to be heterozygous (Supplementary Fig. S3a and b). Subsequently, we performed direct sequencing of this region and found a 3-base insertion (4578-4579insCAT in M33388), that was not previously reported, in the region corresponding to the reverse primer site of the reported Taqman assay in these individuals (Supplementary Fig. S3c). Hence, we performed a copy number analysis by the Taqman assay using a new reverse primer, and confirmed the PCR-RETINA result to be correct (Supplementary Fig. S3d).

Estimation of Allele Copy Number

Through these experiments, we found that PCR-RETINA could estimate the copy number of each allele by combining the information of dot positions in the AD plot graph and the total copy number measured by the Taqman assay. To confirm the accuracy of this estimation, we conducted the standard curve analysis for the rs4756975 locus using CEU individuals. Standard samples were prepared by mixing the two-copy homozygote samples (NA07034 and NA12056) in the range of allelic ratios from 8:1 to 1:8. We performed PCR-RETINA in standard samples with various allelic ratios and analyzed them at 3 minutes of Invader reaction (Fig. 3a). We plotted the log of the fluorescence intensity ratio (FAM intensity/Yellow intensity) on the X-axis, the log of the allele mixing ratio on the Y-axis, and calculated a linear regression curve (Fig. 3b). This standard curve was adjusted using the fluorescence intensity ratio of two-copy heterozygous individuals. After the measurement of the total copy number by the Taqman assay, we performed PCR-RETINA and calculated the copy number of each allele by the linear regression curve. Individuals with various allelic ratios estimated by standard curve analysis were clearly separated by relative dot position in the AD plot (Fig. 3c and d).

Applying PCR-RETINA To Refine Breakpoints of Genomic Duplication/Multiplication

We then applied PCR-RETINA to refine the breakpoints of genomic duplication/multiplication, which is essential to judge whether the functional unit or a part of a certain gene was duplicated or multiplied. From the Database of Genomic Variants, four CNV regions were reported around the *MRGPRX1* gene: Variation_0415 (8 kb) and Variation_0416 (13.4 kb) were found from the fosmid paired-end sequence, Variation_2907 (52.8 kb) from the SNP mapping array, and Variation_3838 (263.3kb) from BAC-array CGH. We designed 26 SNP-based assays covering the largest candidate region (Variation_3838) and performed PCR-RETINA for 90 CEU individuals (Supplementary Fig. S4). Our data clearly indicated that the allelic asymmetries were limited only within the region between rs2220067 and rs7110426, and no additional cluster was found at any loci outside of this region (Fig. 4 and Supplementary Fig. S5). Therefore, we considered that the boundaries of the duplicated/multiplied region were likely to be located between rs12364167 and rs2220067 on one side, and between rs7110426 and rs11024893 on the other side. We confirmed this result by four Taqman assays at the adjacent region of the boundaries. Consequently, we defined only the *MRGPRX1* gene to be present in this CNV region (Fig. 4). Since the University of California Santa Cruz (UCSC) Genome Browser indicated that the boundaries of both sides of the duplicated/multiplied region were located within long interspersed repetitive element 1 (LINE-1) repeats (Fig. 4), we presume that this CNV of *MRGPRX1* occurred within these repeats possibly as a result of nonallelic homologous recombination [Burwinkel and Kilimann, 1998]. The same analysis using YRI individuals also supported the data obtained from CEU individuals (data not shown).

Evaluation of Multiplex PCR RETINA (mPCR-RETINA)

Last, we evaluated the feasibility of the combination of multiplex PCR and RETINA (mPCR-RETINA). We performed 26-plex PCR and RETINA corresponding to the 26 loci mentioned above and compared the AD plot patterns in each locus between 26-plex PCR and PCR with a single primer set. Although we used only 10 ng of genomic DNA for the 26 SNP loci (0.38 ng/site) in the multiplex PCR, the patterns of the AD plots of RETINA were almost identical to those of the single fragments (Supplementary Fig. S5). This suggests that multiplex PCR products can be sufficiently applied as a template of RETINA. Since we have been successfully performing 96-plex PCR using 10 ng of genomic DNA for SNP genotyping, the amount of genomic DNA for mPCR-RETINA will be reduced to a minimum of 0.1 ng for one SNP locus.

DISCUSSION

It is crucial to refine the breakpoints of CNVs to judge whether or not CNVs represent the copy number differences in the

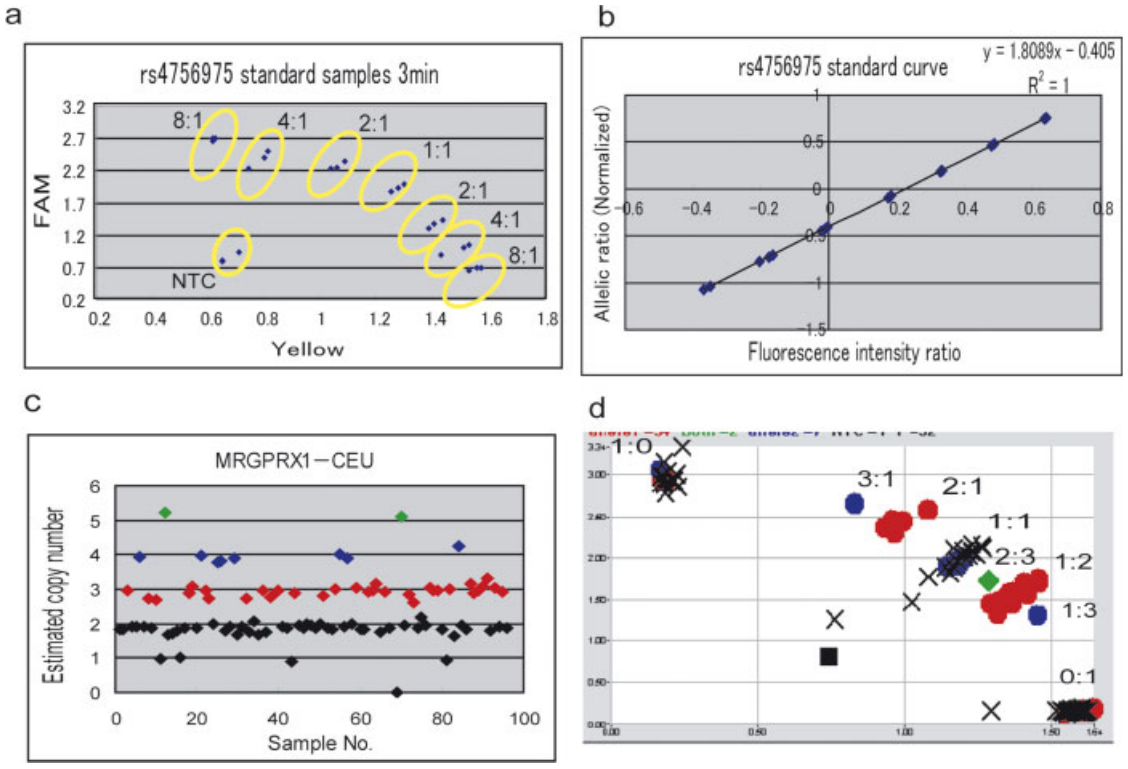


FIGURE 3. **Confirmation of the allele copy number by the standard curve method. a: AD plot of standard samples with various allelic ratios (8:1, 4:1, 2:1, 1:1, 1:2, 1:4, and 1:8) at a 3-minute time-point of the Invader reaction. b: Standard curve by standard samples with known allelic ratio. c: Total copy number obtained by Taqman assay. Closed red diamonds, blue diamonds, and green diamonds indicate three-copy, four-copy, and five-copy individuals, respectively. Closed black diamonds indicate two-copy, one-copy, and zero-copy individuals. d: The AD plot of CEU samples at the 3-minute RETINA reaction. Closed red circles, closed blue circles, and green diamonds indicate individuals with three copies, four copies, and five copies, respectively. X indicates two-copy, one-copy, or zero-copy individuals. Closed black square indicates no template control (NTC). The numbers in the AD plot indicate the estimated allele ratios calculated by the standard curve.**

functional gene unit. Also, it is essential to determine the allele copy number in the CNV regions that may be associated with various phenotypes in medical genetics and pharmacogenetics. Using mPCR-RETINA, we demonstrated here that allelic-asymmetry analysis of a particular locus could define the region of the genomic duplication/multiplication. Additionally, we showed that mPCR-RETINA could accurately determine the allelic ratio and could estimate the allele copy number in CNV regions by a combination with real-time quantitative PCR.

We first examined whether the PCR Invader assay, our standard SNP genotyping method, could be applicable to detect CNVs. Real-time fluorescence monitoring using artificial templates revealed that heterozygote samples with various allelic ratios were clearly separated in an early stage of the reaction, but were merged as one heterozygote cluster in a later phase. This disappearance of allelic asymmetries in the later phase of the Invader reaction is caused by a saturation effect of fluorescence signals originating from a depletion of FRET probes. Real-time fluorescence signal curves showed that both allele signals were saturated in the later phase of the Invader reaction (Fig. 2a). Hence, heterozygote samples with various allelic ratios eventually all come to have the same signal intensity and merge as one heterozygote cluster. Thus, real-time fluorescence monitoring (RETINA) is needed for an efficient detection of allelic asymmetries in the Invader assay.

When we applied PCR-RETINA to the human genome, we found that with PCR-RETINA it is possible to infer the allelic ratio with good precision, but impossible to infer the overall copy number within CNV regions. This characteristic of PCR-RETINA makes it difficult to detect individuals with multiplied allelic symmetries, homozygous individuals with duplications/multiplications, or individuals with deletions. The main reason why PCR-RETINA cannot detect the overall copy number is another saturation effect, namely, a plateau effect of PCR products. In general, a quantity of PCR products reflects the initial copy number of template DNA (genomic DNA) in an exponential phase of the PCR reaction, but does not reflect it in a later phase due to a plateau effect. In our present protocol, most of the PCR products are in the plateau phase and do not reflect the initial copy number of template genomic DNA. In contrast, the allelic ratio is constant through the PCR reaction because the plateau effect of PCR equally influences the amplification efficiencies of both alleles. Additionally, the difference in the input quantity of genomic DNA among the samples has no influence on the allelic ratio, though it usually affects the estimation of the total copy number. As a result, the plateau effect provides sufficient templates with accurate allelic ratios for RETINA and gives good signal intensities within a few minutes of the Invader reaction. For these reasons, PCR-RETINA enables us to estimate the allelic ratio with good precision. To estimate both the total copy number and allele copy number by mPCR-RETINA, further modification is needed.

To our knowledge, up to now melting curve analysis [Ruiz-Ponte et al., 2000; Timmann et al., 2005], the SNP-based real-time PCR-based method using the Taqman assay [Lo et al., 2003; Yu et al.,

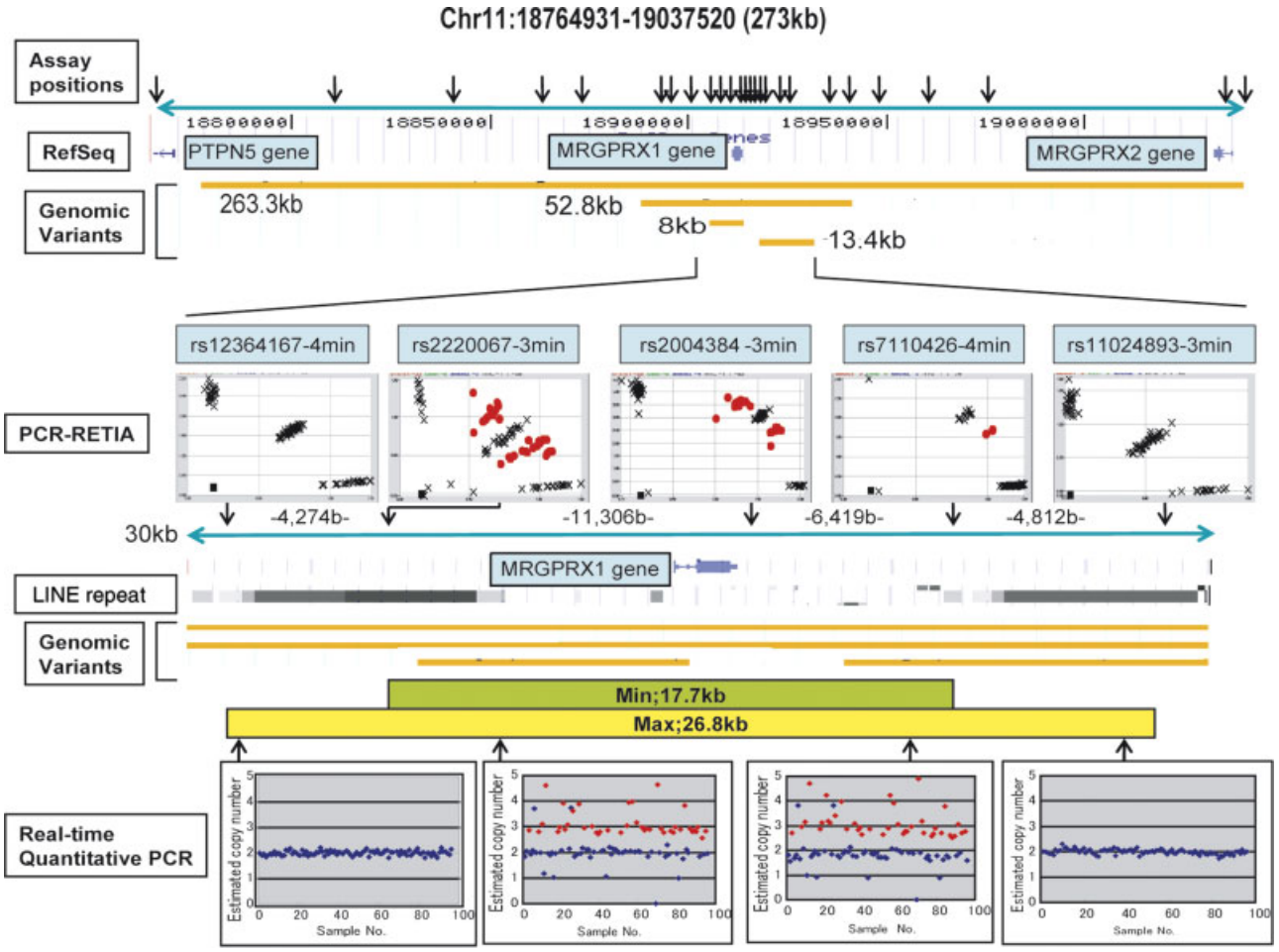


FIGURE 4. **The refinement of the breakpoints of the duplicated/multiplied region including the *MRGPRX1* gene. The region covering the largest genomic variation (Variation_3838) is shown in the upper side of the figure. The map of the *MRGPRX1* gene and its flanking regions are extracted from the UCSC Genome Browser and Database of Genomic Variants. Arrows indicate assay positions of PCR-RETINA and Taqman assays. Representative AD plots especially related to the determination of boundaries of the duplicated/multiplied region are shown in the center of the figure. Closed red circles in the AD plots indicate the individuals with allelic asymmetries. Copy number measurements by the Taqman assays are shown in the lower side of the figure. Red closed diamonds indicate the individuals with allelic asymmetries in at least one of the 26 assays. The boxes marked Min and Max indicate the minimum and maximum size of the duplicated/multiplied regions refined by PCR-RETINA.**

2006], and the single-base extension on matrix-assisted laser-desorption/ionization–time-of-flight (MALDI-TOF) mass spectrometry [Tsui et al., 2005] have been used to detect allelic asymmetries or measure the allelic ratio. However, melting curve analysis and the Taqman assay require expensive target-specific fluorescence probes and a large amount of genomic DNA, and are not suitable for multiple target sites. The single-base extension on MALDI-TOF mass spectrometry involves multiple steps and is labor-intensive. In contrast, PCR-RETINA needs only PCR and the Invader assay, and is very simple and rapid. The adoption of multiplex PCR and no need to synthesize target-specific fluorescence probes can minimize assay costs and save genomic DNA [Lyamichev et al., 1999; Ohnishi et al., 2001]. Additionally, mPCR-RETINA has an advantage in target-specific reactions even in regions of repetitive sequences. The design flexibility of amplicon size (0.1 kb–1 kb or longer) in multiplex PCR facilitates finding unique sequences for PCR primers. Also, the characteristic of the Cleavase VIII enzyme (Third Wave Technologies) that recognizes the specific triplet structure at a target SNP site provides a much higher target-specific reaction compared to the methods by hybridization alone [Lyamichev et al., 1999]. As a proof of that, PCR-RETINA successfully performed three SNP assays (rs2220067, rs11517776 and rs11024893) within Line-1 repeats (Fig. 4; Supplementary Fig. S5).

As PCR-RETINA is a post-PCR detection method, there is a possible risk of cross-contamination. To prevent or minimize this risk, we physically segregate the sample setup and the postreaction workup in the laboratory. In addition, most of the liquid handling processes are conducted by credible robotics in the cleanest environment possible and filtered tips are used in manual mixing or dispensing. We always have "no template control" wells in every plate to monitor cross-contamination and great care is taken to prevent such contamination.

In conclusion, we developed mPCR-RETINA as a new method for detecting allelic asymmetries in CNV regions. mPCR-RETINA can be used for a refinement of duplicated/multiplied regions and the measurement of the copy number of each allele in CNV regions by combining it with real-time quantitative PCR. mPCR-RETINA has clear advantages in saving genomic DNA, target-specific reactions, ease of use, and cost effectiveness. Although

many CNV regions have been reported so far, the functional significance in most of these regions has not been clarified yet. By using mPCR-RETINA and real-time quantitative PCR, we can determine the copy number of each allele in multiple SNP loci of CNV regions efficiently, and examine the relationship of CNV regions with the pathogenesis of common diseases and drug responses. We believe mPCR-RETINA will be a powerful tool in elucidating the functional significance of CNVs for various critical phenotypes.

ACKNOWLEDGMENTS

We thank Mamoru Kato and Tatsuhiko Tsunoda for helpful suggestions in CNV analysis; Rumi Katashima and Tomonaga Matsushita for excellent technical assistance; Nobuhiro Eto and Katsunori Tanaka for data analysis; Tetsuo Mikuriya for the management of oligonucleotides; and Taisei Mushiroda and Aritoshi Iida for suggestions regarding *CYP2D6* assays.

REFERENCES

Aitman TJ, Dong R, Vyse TJ, Norsworthy PJ, Johnson MD, Smith J, Mangion J, Roberton-Lowe C, Marshall AJ, Petretto E, Hodges MD, Bhangal G, Patel SG, Sheehan-Rooney K, Duda M, Cook PR, Evans DJ, Domin J, Flint J, Boyle JJ, Pusey CD, Cook HT. 2006. Copy number polymorphism in *Fcgr3* predisposes to glomerulonephritis in rats and humans. Nature 16:851–855.

Bodin L, Beaune PH, Loriot MA. 2005. Determination of cytochrome P450 2D6 (*CYP2D6*) gene copy number by real-time quantitative PCR. J Biomed Biotechnol 3:248–253.

Burwinkel B, Kilimann MW. 1998. Unequal homologous recombination between LINE-1 elements as a mutational mechanism in human genetic disease. J Mol Biol 3:513–517.

Dorado P, Caceres MC, Pozo-Guisado E, Wong ML, Licinio J, Llerena A. 2005. Development of a PCR-based strategy for *CYP2D6* genotyping including gene multiplication of worldwide potential use. Biotechniques 39:571–574.

Freeman JL, Perry GH, Feuk L, Redon R, McCarroll SA, Altshuler DM, Aburatani H, Jones KW, Tyler-Smith C, Hurles ME, Carter NP, Scherer SW, Lee C. 2006. Copy number variation: new insights in genome diversity. Genome Res 16:949–961.

Gasche Y, Daali Y, Fathi M, Chiappe A, Cottini S, Dayer P, Desmeules J. 2004. Codeine intoxication associated with ultrarapid *CYP2D6* metabolism. N Engl J Med 30:2827–3281.

Gonzalez E, Kulkarni H, Bolivar H, Mangano A, Sanchez R, Catano G, Nibbs RJ, Freedman BI, Quinones MP, Bamshad MJ, Murthy KK, Rovin BH, Bradley W, Clark RA, Anderson SA, O'Connell RJ, Agan BK, Ahuja SS, Bologna R, Sen L, Dolan MJ, Ahuja SK. 2005. The influence of *CCL3L1* gene-containing segmental duplications on HIV-1/AIDS susceptibility. Science 4:1434–1440.

Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, Scherer SW, Lee C. 2004. Detection of large-scale variation in the human genome. Nat Genet 36:949–951.

International HapMap Consortium. 2005. A haplotype map of the human genome. Nature 27:1299–1320.

Locke DP, Sharp AJ, McCarroll SA, McGrath SD, Newman TL, Cheng Z, Schwartz S, Albertson DG, Pinkel D, Altshuler DM, Eichler EE. 2006. Linkage disequilibrium and heritability of copy-number polymorphisms within duplicated regions of the human genome. Am J Hum Genet 79:275–290.

Lo HS, Wang Z, Hu Y, Yang HH, Gere S, Buetow KH, Lee MP. 2003. Allelic variation in gene expression is common in the human genome. Genome Res 13:1855–1862.

Lovlie R, Daly AK, Molven A, Idle JR, Steen VM. 1996. Ultrarapid metabolizers of debrisoquine: characterization and PCR-based detection of alleles with duplication of the *CYP2D6* gene. FEBS Lett 19:30–34.

Lyamichev V, Mast AL, Hall JG, Prudent JR, Kaiser MW, Takova T, Kwiatkowski RW, Sander TJ, de Arruda M, Arco DA, Neri BP, Brow MA. 1999. Polymorphism identification and quantitative detection of genomic DNA by invasive cleavage of oligonucleotide probes. Nat Biotechnol 17:292–296.

Mast A, de Arruda M. 2006. Invader assay for single-nucleotide polymorphism genotyping and gene copy number evaluation. Methods Mol Biol 335:173–186.

Nevilie M, Selzer R, Aizenstein B, Maguire M, Hogan K, Walton R, Welsh K, Neri B, de Arruda M. 2002. Characterization of cytochrome P450 2D6 alleles using the Invader system. Biotechniques 32:S34–S43.

Ohnishi Y, Tanaka T, Ozaki K, Yamada R, Suzuki H, Nakamura Y. 2001. A high-throughput SNP typing system for genome-wide association studies. J Hum Genet 46:471–477.

Padiath QS, Saigoh K, Schiffmann R, Asahara H, Yamada T, Koeppen A, Hogan K, Ptacek LJ, Fu YH. 2006. Lamin B1 duplications cause autosomal dominant leukodystrophy. Nat Genet 38:1114–1123.

Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, Cho EK, Dallaire S, Freeman JL, Gonzalez JR, Gratacos M, Huang J, Kalaitzopoulos D, Komura D, MacDonald JR, Marshall CR, Mei R, Montgomery L, Nishimura K, Okamura K, Shen F, Somerville MJ, Tchinda J, Valsesia A, Woodwark C, Yang F, Zhang J, Zerjal T, Armengol L, Conrad DF, Estivill X, Tyler-Smith C, Carter NP, Aburatani H, Lee C, Jones KW, Scherer SW, Hurles ME. 2006. Global variation in copy number in the human genome. Nature 23:444–454.

Ruiz-Ponte C, Loidi L, Vega A, Carracedo A, Barros F. 2000. Rapid real-time fluorescent PCR gene dosage test for the diagnosis of DNA duplications and deletions. Clin Chem 46:1574–1582.

Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, Maner S, Massa H, Walker M, Chi M, Navin N, Lucito R, Healy J, Hicks J, Ye K, Reiner A, Gilliam TC, Trask B, Patterson N, Zetterberg A, Wigler M. 2004. Large-scale copy number polymorphism in the human genome. Science 23:525–528.

Sharp AJ, Locke DP, McGrath SD, Cheng Z, Bailey JA, Vallente RU, Pertz LM, Clark RA, Schwartz S, Segraves R, Oseroff VV, Albertson DG, Pinkel D, Eichler EE. 2005. Segmental duplications and copy-number variation in the human genome. Am J Hum Genet 77:78–88.

Sharp AJ, Cheng Z, Eichler EE. 2006. Structural variation of the human genome. Annu Rev Genomics Hum Genet 7:407–442.

Timmann C, Moenkemeyer F, Evans JA, Foerster B, Tannich E, Haase S, Sievertsen J, Kohne E, Horstmann RD. 2005. Diagnosis of alpha+-thalassemias by determining the ratio of the two alpha-globin gene copies by oligonucleotide hybridization and melting curve analysis. Clin Chem 51:1711–1713.

Tsui NB, Chiu RW, Ding C, El-Sheikhah A, Leung TN, Lau TK, Nicolaides KH, Lo YM. 2005. Detection of trisomy 21 by quantitative mass spectrometric analysis of single-nucleotide polymorphisms. Clin Chem 51:2358–2362.

Tuzun E, Sharp AJ, Bailey JA, Kaul R, Morrison VA, Pertz LM, Haugen E, Hayden H, Albertson D, Pinkel D, Olson MV, Eichler EE. 2005. Fine-scale structural variation of the human genome. Nat Genet 37:727–732.

Urban AE, Korbel JO, Selzer R, Richmond T, Hacker A, Popescu GV, Cubells JF, Green R, Emanuel BS, Gerstein MB, Weissman SM, Snyder M. 2006. High-resolution mapping of DNA copy alterations in human chromosome 22 using high-density tiling oligonucleotide arrays. Proc Natl Acad Sci USA 21:4534–4539.

Yu A, Geng H, Zhou X. 2006. Quantify single nucleotide polymorphism (SNP) ratio in pooled DNA based on normalized fluorescence real-time PCR. BMC Genomics 7:143.